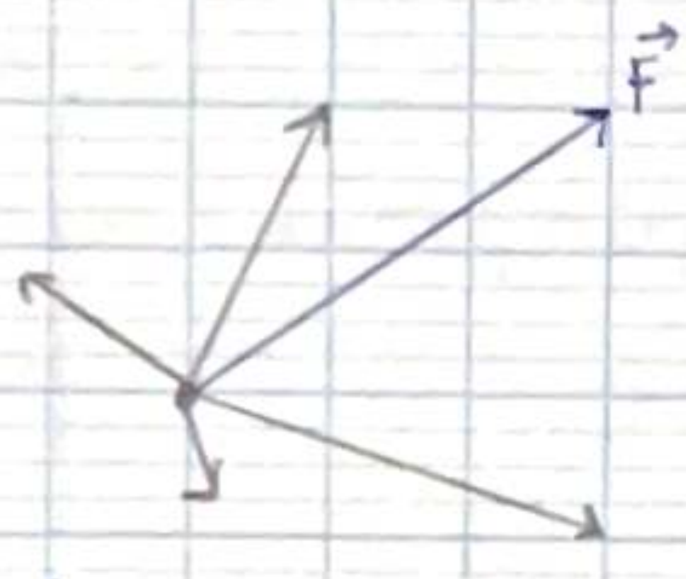
Einstein Field Equations Derivation.

The Principle of least Action. & Calculus of Variation.

Newtonian Mechanics

Consider a single particle at position $\vec{r}(t)$ acted upon by a force $\vec{F}$. This force might be a resultant of some different forces.



Newton's Second law states that,

$$\vec{F} = m\vec{a} = m\ddot{\vec{r}}$$

For conservative forces (gravity, electrostatics, but not friction), the force can be expressed in terms of a potential.

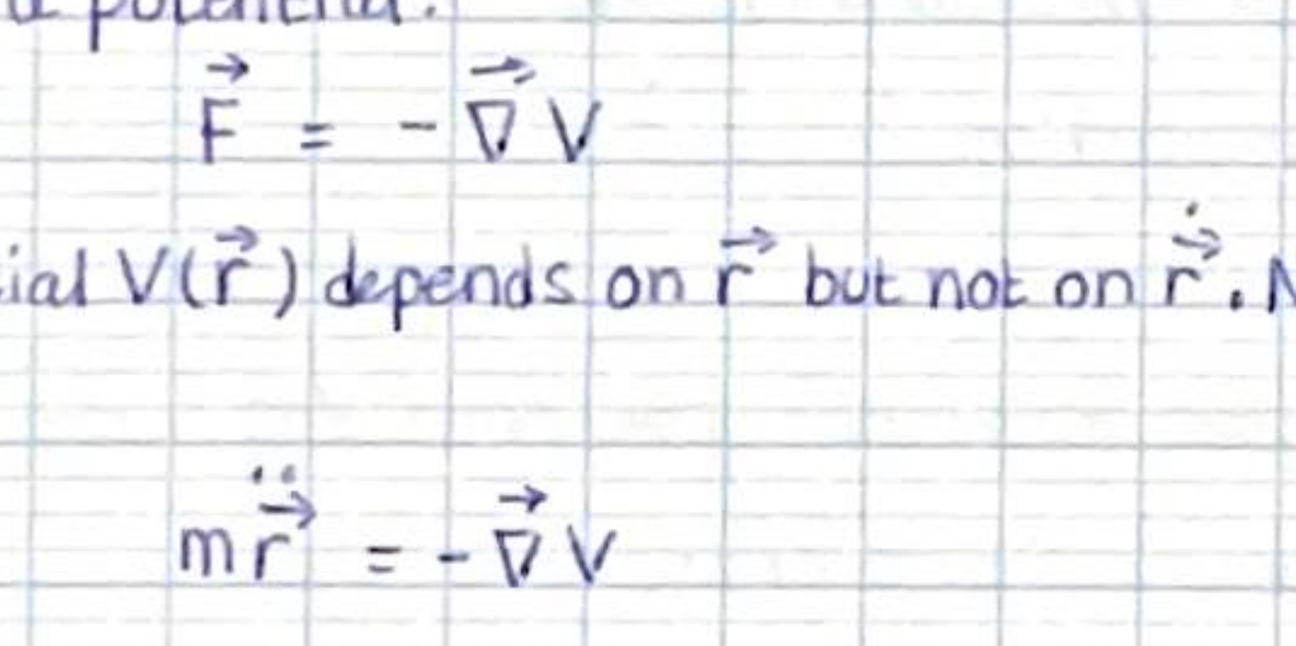
$$\vec{F} = -\vec{\nabla} V$$

The potential $V(\vec{r})$ depends on $\vec{r}$ but not on $\dot{\vec{r}}$. Newton's equation then reads

$$m\ddot{\vec{r}} = -\vec{\nabla} V$$

This is a second order differential equation whose general solution has two integration constants. Physically, this means we have to specify the initial ~~and final positions, $\vec{r}(t_1)$ and $\vec{r}(t_2)$~~ position $\vec{r}(t_1)$ and velocity $\dot{\vec{r}}(t_1)$ of the particle to figure out where it is going to end up.

Instead of specifying initial position and velocity, let us specify initial position $\vec{r}(t_1)$ and final position $\vec{r}(t_2)$ and consider the possible paths that connect them.



What path does the particle actually take?

$$\cancel{E_{Tot}} = \int_{t_1}^{t_2}$$

$$\underbrace{E(\vec{r}(t))}_{E_{tot}} = \int_{t_1}^{t_2} (E_k + E_p)\,dt$$

+ for Total energy -

To each path $\vec{r}(t)$ let a number (called the action be assigned).

$$S(\vec{r}(t)) = \int_{t_1}^{t_2} dt \left(\tfrac{1}{2} m \dot{\vec{r}}^{\,2} - V(\vec{r})\right)$$

For Action : $(E_K - E_P)$

For total energy : $(E_K + E_P)$

$$S = \int_{t_1}^{t_2} L \, dt$$

where L is the lagrangian which is defined as

$$L = (KE - PE)$$

For the action integral to be well-defined, the trajectory has to be bounded in space and time.

$$\boxed{\text{Note: The true path taken by the particle is an extremum of action } S.}$$

Proof

To find the extremum of a function, we differentiate it and set it equal to zero. But the action is not a function; It is a functional - a function of a function. This makes it a slightly different problem.

S is action which is a functional ie. it takes a function of time (for fields) space as input and returns a scalar.

Consider a given path $\vec{r}(t)$. We ask how the action changes when we change the path slightly

$$\vec{r}(t) \to \vec{r}(t) + \delta\vec{r}(t)$$

but in such a way that we keep the end points of the path fixed

$$\delta\vec{r}(t_1) - \delta\vec{r}(t_2) = 0$$

The action for the perturbed path $\vec{r} + \delta\vec{r}$ is

$$S(\vec{r} + \delta\vec{r}) = \int_{t_1}^{t_2} dt \left(\frac{1}{2} m\left(\dot{\vec{r}}^{\,2} + 2\dot{\vec{r}} \cdot \delta\dot{\vec{r}} + \delta\dot{\vec{r}}^{\,2}\right) - V(\vec{r} + \delta\vec{r})\right)$$

The potential can be Taylor expanded

$$V(\vec{r} + \delta\vec{r}) = V(\vec{r}) + \vec{\nabla} V \cdot \delta\vec{r} + O(\delta\vec{r}^{\,2})$$

Gradient/derivative

Big O notation

Describes the limiting behaviour of a function when the argument tends towards a particular value or infinity.

The difference between the action for the perturbed path and the unperturbed path is

$$\delta S \equiv S(\vec{r}+\delta\vec{r}) - S(\vec{r}) = \int_{t_1}^{t_2} dt \left[m\dot{\vec{r}}\cdot\delta\dot{\vec{r}} - \vec{\nabla}\cdot\delta\vec{r} \right] + \dots \quad -①$$

$$= \int_{t_1}^{t_2} \left[-m\ddot{\vec{r}} - \vec{\nabla}V\right]\cdot\delta\vec{r} + \left[m\dot{\vec{r}}\cdot\delta r\right]_{t_1}^{t_2} \quad -②$$

Note: To get from ① to ②, an integration by parts ~~is~~ was carried out. This picks up a term that was evaluated at the boundaries t_1 and t_2. However, this term vanishes since the end points are fixed, $\underbrace{\delta(\vec{r}_1) = \delta(\vec{r}_2)}_{=0}$ & $\delta(\vec{r}_1) - \delta(\vec{r}_2) = 0$

Hence we get,

$$\delta S = \int_{t_1}^{t_2} \left[-m\ddot{\vec{r}} - \vec{\nabla}V\right]\cdot\delta\vec{r}$$

The condition that the path we started with is an extremum of the action is

$$\delta S = 0$$

This should hold for all changes $\delta\vec{r}(t)$ that could be made to the path. The only way this can be true is that the integrand is zero.

$$-m\ddot{\vec{r}} - \vec{\nabla}V = 0$$

$$\Rightarrow \boxed{m\ddot{\vec{r}} = -\vec{\nabla}V}$$

This is Newton's Second law requiring that the action is an extremum is equivalent to requiring that the path obeys Newton's equation.

The integrand of the action is called the Lagrangian $L(\vec{r}, \dot{\vec{r}}, t)$ ie. $S = \int_{t_1}^{t_2} L \, dt$
In our previous example, we have $L = \frac{1}{2} m \dot{\vec{r}}^2 - V(\vec{r})$. From the lagrangian, we can define a generalised momentum and a generalised force.

$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{r}}} \quad \text{and} \quad F = \frac{\partial L}{\partial \vec{r}}$$

$P = m\dot{\vec{r}}$ and $\vec{F} = \vec{\nabla} V$

Newton's equation can be written as

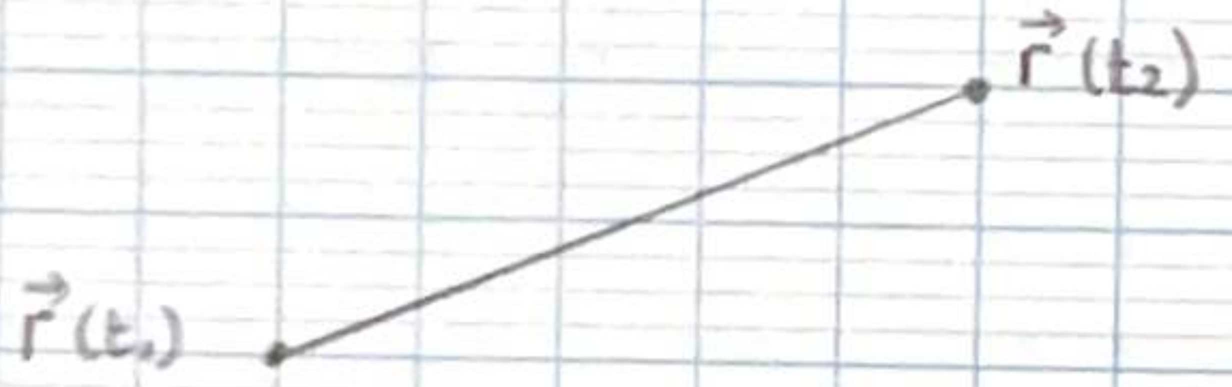
$$\frac{d\vec{p}}{dt} = \vec{F} \iff \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{r}}}\right) = \frac{\partial L}{\partial \vec{r}}$$

This is called the Euler-Lagrange equation.

Consider the lagrangian of a free particle. (1D)

$$L = \frac{1}{2} m \dot{\vec{r}}^2$$

In this case, the least-action principle implies that we want to minimise kinetic energy over a fixed time. So, the particle must take a direct route which is a straight line.



But do we slow down to begin with then speed up? Or, do we accelerate at the beginning then slow down? Or, do we go at uniform speed? The average speed is fixed to be the total distance over total time. So, if anything is done except uniform speed, we are sometimes going too fast or too slowly. But the mean of the square of something (in this case velocity) is that deviates around an average, is always greater than the square of the mean. Hence, to minimise the integrated kinetic energy - ie. the action - the particle should go at uniform speed. In the absence of potential, this is what we should get.

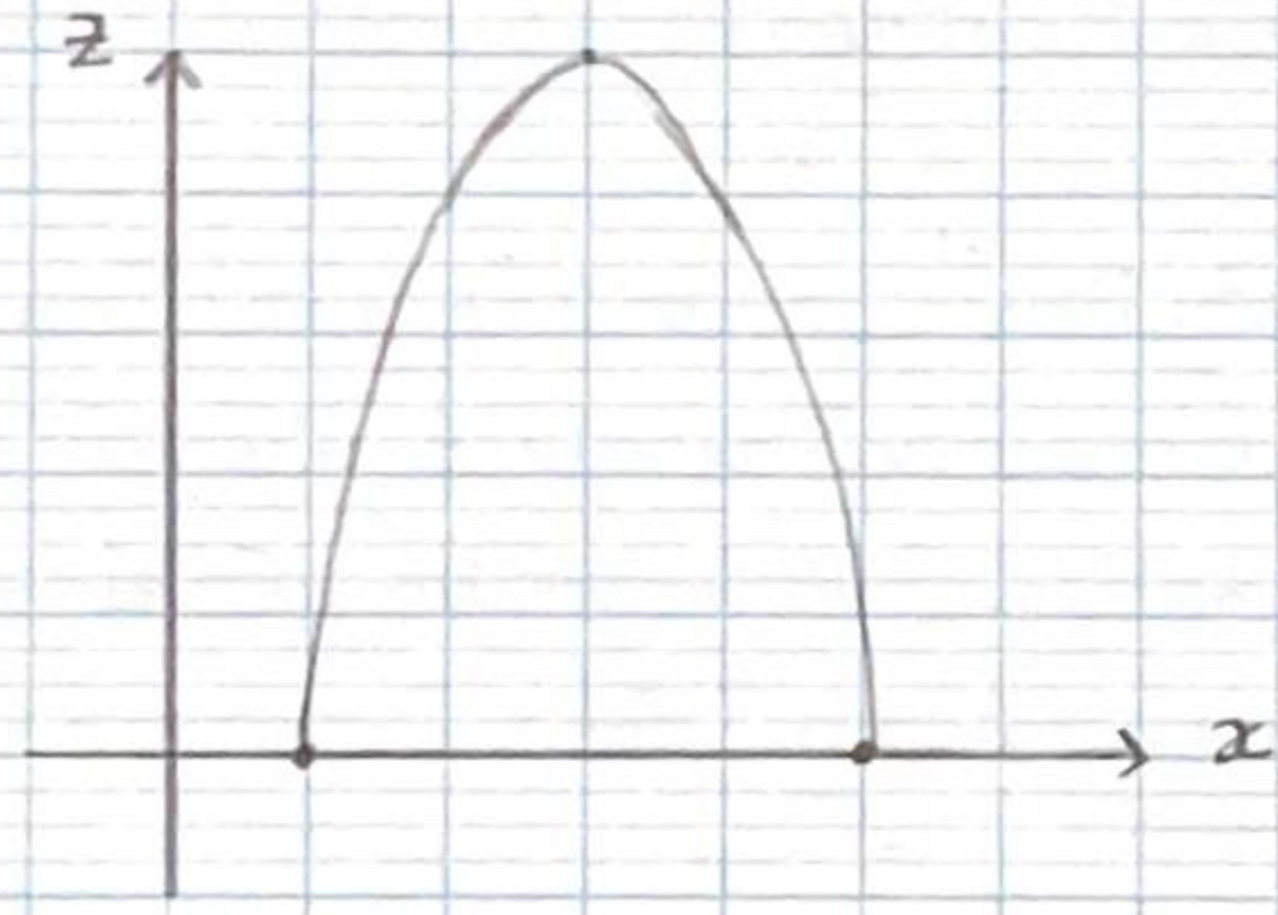
why does a fan produce cool air - Physics

Consider the Lagrangian of a particle in a uniform gravitational field (2D)
The lagrangian is :

$$\cancel{\mathcal{L} = \frac{1}{2} m v^2}$$

$$L = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} m\dot{z}^2 - mgz$$

Imagine that the particle starts and ends at the same height $z_0 = z(t_1) = z(t_2)$, but moves horizontally from $x_1 = x(t_1)$ to $x_2 = x(t_2)$. This time, we don't want to go in a straight line. Instead we can minimise the difference between K·E and P·E ; if we go up where the P·E is bigger. But we don't want to go to high either since that requires a lot of K·E to begin with. To strike the right balance, the particle takes a parabola.



Unification of Physics

The lagrangian method is nowadays used to describe all of physics, not just mechanics, this include electromagnetism, special and general relativity, particle physics and even string theory where all the fundamental laws can be expressed in terms of least action principle. Nearly every experiment performed can be explained by the lagrangian of the Standard model.

$$L \sim \underbrace{R}_{\text{Einstein}} \underbrace{\overbrace{- \frac{1}{4} F_{\mu\nu} F^{\mu\nu}}^{\text{Maxwell}}}_{\text{Yang-Mills}} + \underbrace{i\bar{\Psi}\gamma^{\mu} D_{\mu}\Psi}_{\text{Dirac}} + \underbrace{|D_{\mu}h|^2 - V(|h|)}_{\text{Higgs}} + \underbrace{h\bar{\Psi}\Psi}_{\text{Yukawa}}.$$

The first 2 terms characterise all fundamental forces in Nature : The term Einstein describes gravity. Blackholes and the expansion of the universe follow from it. The next term Maxwell describe electric and magnetic forces which are just different manifestations of a single(electromagnetic) force. A generalisation of this (Yang - Mills) encodes the strong and weak nuclear forces. The next term Dirac describe all matter particles collectively called fermions things like electrons, quarks and neutrinos. Without the final 2 terms, Higgs and Yukawa, the particles would be massless.

Feynman's Path Integral

According to quantum mechanics, a free particle does not take only one path but all possible paths. The key insight on quantum mechanics is ~~on~~ that nature is probabilistic and things happen by random chance.

The probability that a particle starting at $\vec{r}(t_1)$ and will end up at $\vec{r}(t_2)$ is expressed in terms of an amplitude A, which is a complex number which can be thought as the square root of the probability.

A - Quantum wave function Ψ ??

$$\text{Prob} = |A|^2$$

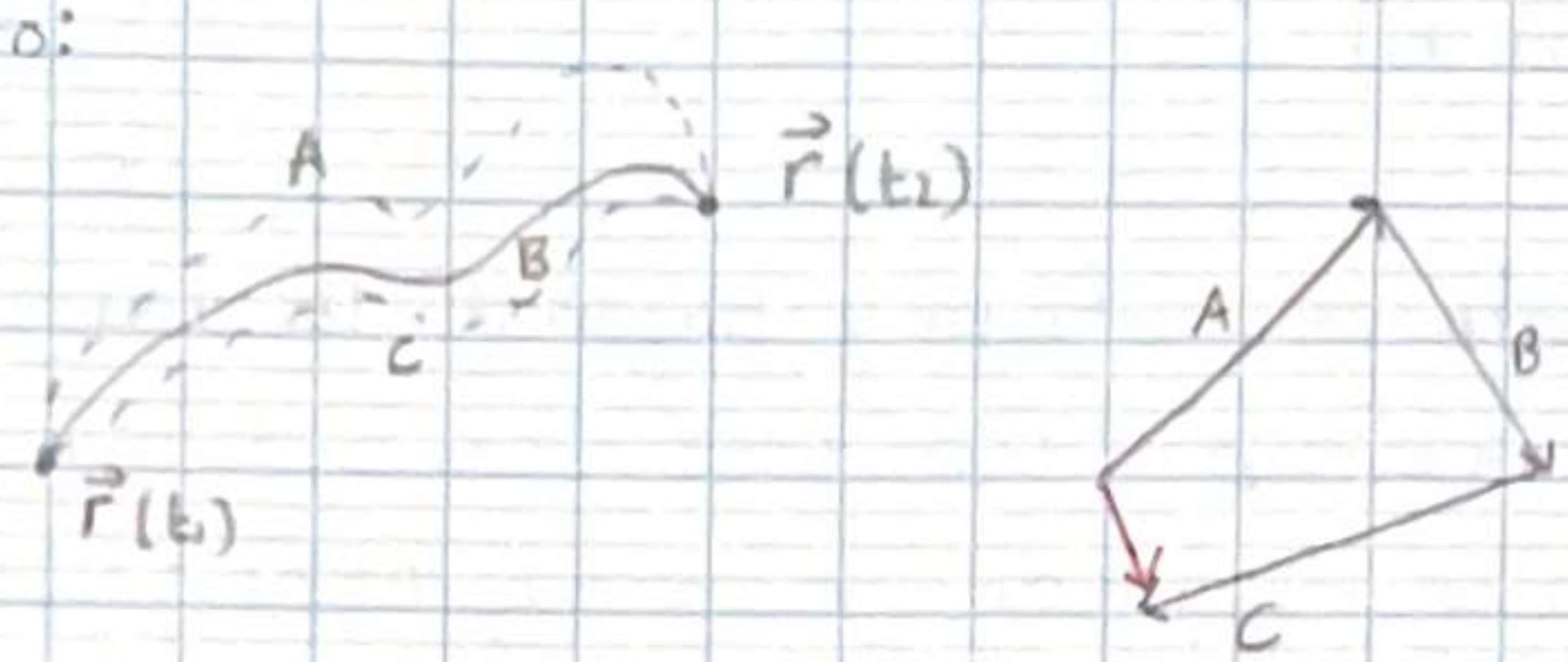
To compute the amplitude, we must sum over all path that the particle can take weighted by a phase.

$$A = \sum_{\text{paths}} e^{iS/\hbar} \quad -(1)$$

Here, the phase for each path is just the value of the action for the path in units of Planck's constant $\hbar = 1.05 \times 10^{-34}$ J.s. Recall that complex numbers can be represented as arrows in the xy-plane. The length of the arrow represents the magnitude of the complex number, the phase represents its angle relative to the x-axis. Here in (1) each path can be represented by an arrow with a phase angle of $S/\hbar$. Summing all arrows then squaring gives the probability. This is Feynman's path integral approach to quantum mechanics sometimes called Sum over histories.

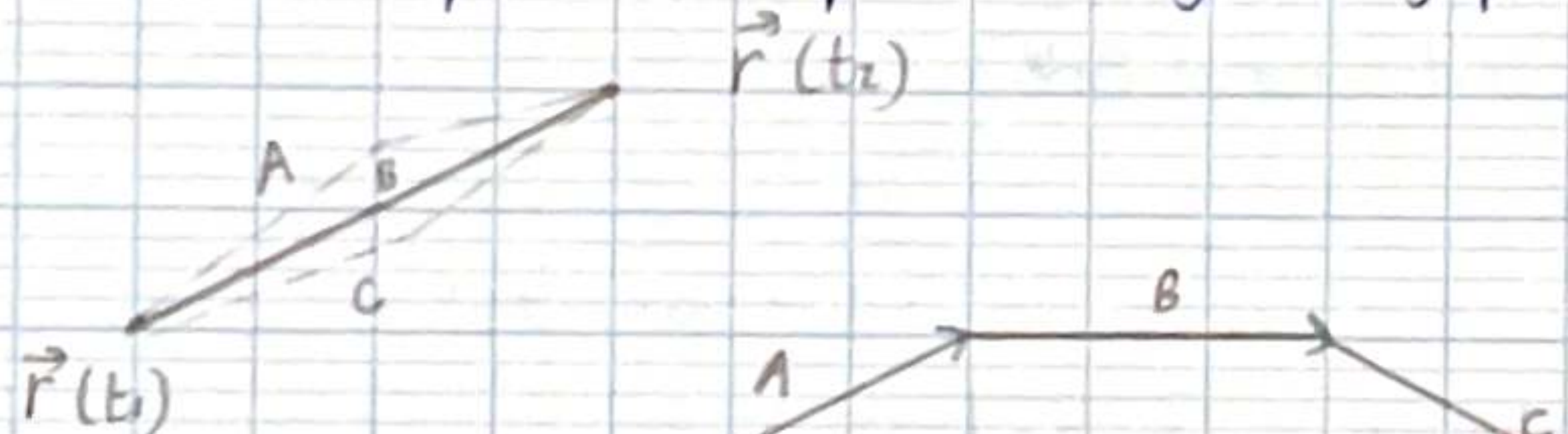
Prove Projectile motion is least action.
↳ straight line is least action.

The way to think about it is as follows: When a particle moves, it really does take all possible paths. However, away from the classical path (the straight line) the action varies wildly between neighbouring paths and the sum of different paths therefore averages to zero:



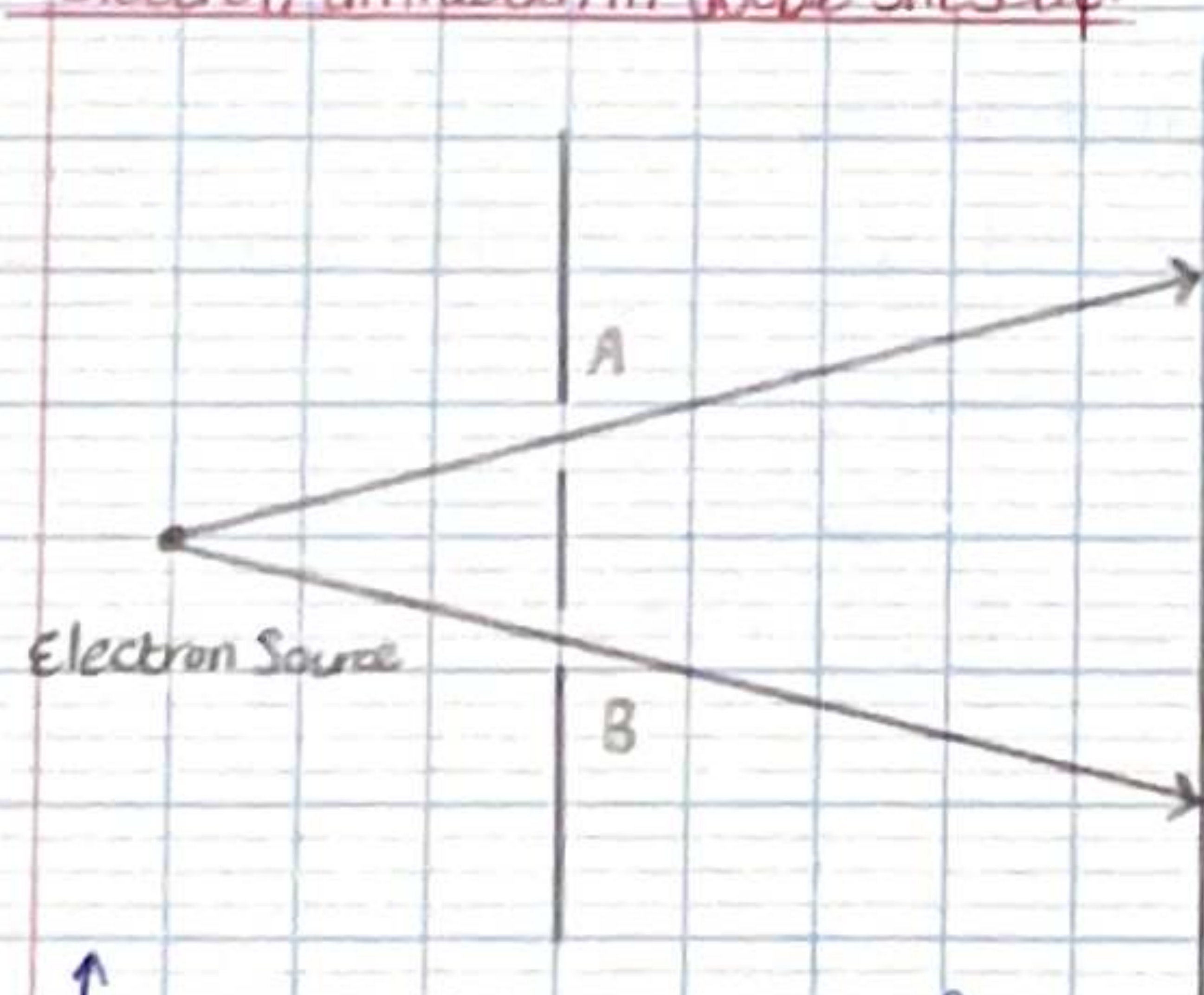
ie. far from the classical path, the little arrows representing the complex amplitudes of each point in random directions and cancel each other.

Only near the classical path do the phase of neighbouring paths reinforce each other:



ie. near the classical path, the little arrows representing the complex amplitudes of each path point in the same direction and therefore add. This is how classical physics emerges from quantum physics

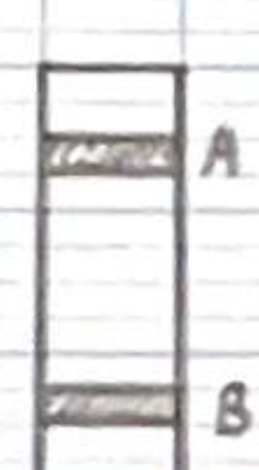
Electron diffraction in double slit setup.



↑ Consider this setup which fires electrons one at a time.

→ In Newtonian physics, each electron either take path A or B. We then expect the image behind the two slits just to be the sum of electrons going through slit A or B; ie. we expect the electrons to build up a pair of stripes - one corresponding to electrons through A, the other through B.

Expected:



In the experiment, we just don't see 2 stripes but a series of stripes. In quantum physics, each electron seem to take path A and B simultaneously and interfere with itself. (Since the electrons are fired one by one, there is nobody else around to interfere with it.)

Find the extremal function $y = f(x)$, which is the shortest curve that connects two points (x_1, y_1) and (x_2, y_2)

The arc length of the curve is given by $A[y] = \int_{x_1}^{x_2} \sqrt{1 + [y'(x)]^2}\, dx$

functional - square brackets !!

with $y' = \frac{dy}{dx}$, $y_1 = f(x_1)$, $y_2 = f(x_2)$

The Euler-Lagrange equation can be used to find the extremal function $f(x)$ that minimises the functional $A[y]$

$$\frac{\partial L}{\partial f} - \frac{d}{dx}\frac{\partial L}{\partial f'} = 0$$

With $L = \sqrt{1 + [f'(x)]^2}$.

Since f does not appear explicitly in L, the first term in the Euler-Lagrange equation vanishes for all $f(x)$ and thus,

$$\frac{d}{dx}\frac{\partial L}{\partial f'} = 0$$

Substituting for L and taking the derivative,

$$\frac{d}{dx}\frac{f'(x)}{\sqrt{1+[f'(x)]^2}} = 0 \quad \text{thus} \quad \frac{f'(x)}{\sqrt{1+[f'(x)]^2}} = c \quad \text{for some constant } c$$

Then $$\frac{[f'(x)]^2}{1+[f'(x)]^2} = c^2 \quad \text{where } 0 \leq c^2 \leq 1$$

Solving, we get

$$[f'(x)]^2 = \frac{c^2}{1-c^2}$$

which implies that

$f'(x) = m$

is a constant and therefore the shortest curve that connects two points (x_1, y_1) and (x_2, y_2) is

$$f(x) = mx + b \text{ with } m = \frac{y_2 - y_1}{x_2 - x_1} \text{ and } b = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}$$

and we have thus found the extremal function $f(x)$ that minimises the functional $A[y]$ so that $A[f]$ is a minimum. Note that $y = f(x)$ is the equation for a straight line and in other words, the shortest distance between 2 points is a straight line.

(2) → Projectile Motion & Calculus of Variations.

Qualitatively explaining the principle of least action in quantum mechanics and diffraction.

Due to uncertainty in quantum mechanics, there is no unique path. However, we can associate each possible path with a definite amount of action.

However instead of taking velocities, we usually take momentum.

- Photons have no rest mass ($m_0 = 0$) but they do have momentum because of their energy.

$E = mc^2$ can be rewritten as $E = P c$ & Einstein Planck relation tells us $E = hf$, for a photon, wavelength is given by $\lambda = \frac{c}{f}$

Hence, $p = E/c = hf/c = h/\lambda = \hbar \cdot k$

$E = mc^2 \rightarrow E_k - E_p = mv^2$

The energy of a photon does not vary (Quantum Theory)

So the value of the integral is ~~t~~ just energy times the travel time.

What is the travel time?

Let r_{12} be defined as a vector pointing from r_1 to r_2

The distance between the 2 points is $|r_{12}| = \sqrt{r_{12}^2} = r_{12}$

Photons travel at the speed of light so time interval $\rightarrow t = r_{12}/c$

$\Rightarrow S = p \cdot c \cdot t = p \cdot c \cdot r_{12}/c = p \cdot r_{12}$

If it is assumed that the direction of momentum of the photon & that of r_{12} is different, a dot product is used

$\Rightarrow S = p \cdot r_{12} = |p|\,|r_{12}| \cdot \cos\theta = P \cdot r_{12} \cdot \cos\theta$ with θ = angle between p & r_{12}.

If the angle is the same, $= 0$, then $\cos\theta = 1$. If the angle is $\pm \pi/2$, then it is 0

Due to dealing with Quantum mechanics, we cannot minimise the action to determine the actual path.

We'll use $S = p \cdot r_{12}$ to calculate a probability amplitude
We'll associate trajectories with amplitude and just use the action values to do so.

① We measure in units of $\hbar$, So, we write $\theta = S/\hbar = p \cdot r_{12}/\hbar$

② θ denotes an angle. $\theta = p \cdot r_{12}/\hbar$ is called the phase of a proper wave function.

$$\hookrightarrow \Psi(p, r_{12}) = a \cdot e^{i \cdot \theta} = (1/r_{12})\, e^{i \cdot p \cdot r_{12}/\hbar}$$

propagator wave function for a photon

→ Usually written using more specific notations as:

$$\langle r_2 | r_1 \rangle = \frac{e^{i p \cdot r_{12}/\hbar}}{r_{12}}$$

This notation is better (looks) because it uses Dirac's bra-ket notation: the initial state of the photon is written as $\langle r_1|$ and final state $|r_2\rangle$
It is the amplitude of the photon to go from point a to b (r_1 to r_2).
the $1/r_{12}$ coefficient is present to cater for the inverse square law.

As everything decreases with distance

We get the inverse square $\left(1/r_{12}^2\right)$ when calculating a probability which is done by taking the absolute square of the amplitude

$$\left|(1/r_{12}) \cdot e^{i \cdot p \cdot r_{12}/\hbar}\right|^2 = \left|1/r_{12}^2\right|^2 \cdot \left|e^{i \cdot p \cdot r_{12}/\hbar}\right|^2 = 1/r_{12}^2$$

Suppose Feynman's Summary on his path integral formulation of quantum mechanics.

Suppose for all paths, S is very large compared to $\hbar$. One path contributes to a certain amplitude. For a nearby path, the phase is quite different, because ~~for~~ with a very large S, even a small change in S means a completely different phase because $\hbar$ is so small. So, nearby paths will normally cancel their effects out in taking the sum, except for one region, and that is when a path and a nearby path all give the same phase in the first approximation (the same action within $\hbar$). Only those paths are considered important.

Formula for the amplitude of photon to go from one point to another

$$\langle r_1 | r_2 \rangle = \frac{e^{i \cdot p \cdot r_{12}/\hbar}}{r_{12}}$$

→ Note: A photon also has an amplitude to travel faster or slower than c.

The photon propagator function can be rewritten as:

$$\psi(p, r_{12}) = a \cdot e^{i\theta} = \frac{1}{r_{12}} e^{i \cdot S/\hbar} = e^{i \cdot p \cdot r_{12}/\hbar} / r_{12}$$

$S = p \cdot r_{12}$ is the amount of action we calculated for the path
Action is energy over time ($1 N \cdot m \cdot s$) = ($1 J \cdot s$) or momentum over some distance.

In quantum mechanics, we have to add the amplitude for the various paths between r_1 and r_2 so we get a 'final arrow' whose absolute square gives us the probability of the photon going from r_1 to r_2. We don't know what actually happens in between but we know that the amplitudes interfere.

From Feynman's book,

Quantum theory says that light can go from a source in air to a detector in water in many ways. ~~The~~

The assumption that light travels at the speed of light reduces the least action principle to least time principle. The arrows which point in the same direction make the biggest contribution to the final arrow.

↳ use stopwatch as metaphor : If stopwatches arrive at the same time, they point in more or less the same direction.

Diffraction.

When we have a large slit, the photon is likely to travel in a straight line. There are many other possible paths - curved / crooked paths but the amplitude associated with those other paths cancel out each other. In contrast, the straight line paths and ~~most~~ importantly, the nearby paths reinforce each other as they are associated with amplitudes having the same phase, more or less.

When the slit is narrow, there are not enough arrows to cancel out each other and therefore the crooked paths also become associated with sizable probabilities.

The phase of our amplitude θ is measured in units of $\hbar$: $\theta = S/\hbar$, Hence, we should measure the variations in S in units of $\hbar$.

Consider 2 paths: one for which the action is equal to S and one for which the action is $S + \delta S = S + \pi\hbar$ so $\delta S = \pi\hbar$ (phase difference of π)

↳ Destructive interference occurs when the phase difference is an odd multiple of π.

They will cancel out each other.

$$e^{i.S/\hbar}/r_{12} + e^{i.(S+\delta S)/\hbar}/r_{12}$$

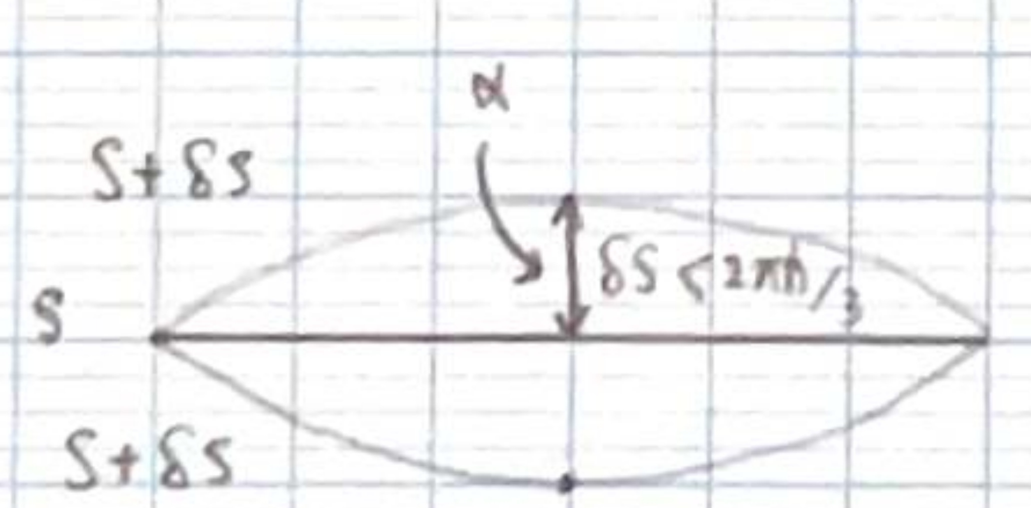
$$= \frac{1}{r_{12}}\left(e^{i\,S/\hbar} + e^{i(S+\pi\hbar/\hbar)}\right)$$

$$= \frac{1}{r_{12}}\left(e^{i.S/\hbar} + e^{i.S/\hbar}\cdot \underbrace{e^{i\pi}}_{-1}\right)$$

$$= \frac{1}{r_{12}}\left(e^{i.S/\hbar} - e^{i.S/\hbar}\right)$$

$$= 0$$

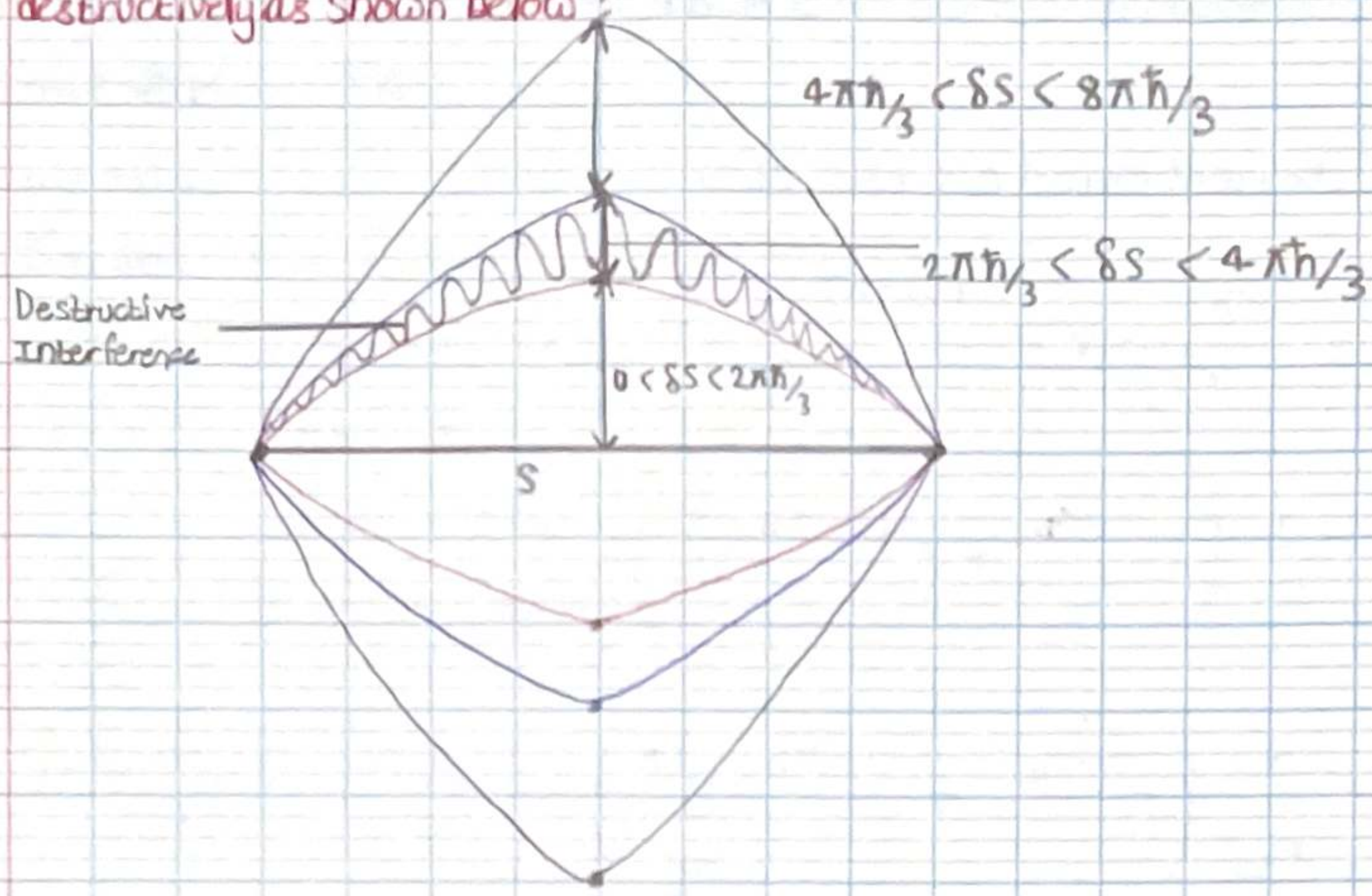
So (very) nearby paths will interfere constructively, making the final arrows larger.
Paths with phase difference of π will interfere destructively.
↳ In order for that to happen, $S\hbar$ should be smaller than $\frac{2}{3}\pi\hbar \approx 2\hbar$



If arrow α is the amplitude which we are adding together, any amplitude whose phase angle is smaller than $2/3\,\pi\hbar \approx 2\hbar$ will add something to its length (increasing probability thus constructive interference).

Note: The amplitudes of the higher and lower path in the drawing do not cancel. On the contrary, the action S is the same, so their magnitudes just add up.
↳ (equiprobable)

There will be alternating zones with paths that interfere constructively and destructively as shown below :



Einstein's Equivalence Principle

Physically, this principle postulates that there is no experiment done in a Small confined space which can tell the difference between a uniform gravitational field and an equivalent accleration.
Mathematically, this observation states that in the presence of a gravitational field, Small enough free falling frames will be inertial and that in these Local Inertial Frames where the metric $g_{\mu\nu}$ reduces to $\eta_{\mu\nu}$, the laws of physics from Special relativity will hold true.